Virology 441 (2013) 85–94

Contents lists available at SciVerse ScienceDirect

Virology

journal homepage: www.elsevier.com/locate/yviro

Detection and initial characterization of protein entities consisting of the HIV glycoprotein cytoplasmic C-terminal domain alone

Tanya Pfeiffer [a], Thomas Ruppert [b], Heiner Schaal [c], Valerie Bosch [a,*]

[a] Forschungsschwerpunkt Infektion und Krebs, Deutsches Krebsforschungszentrum, 69120 Heidelberg, Germany
[b] Zentrum für Molekularbiologie, Universität Heidelberg, 69120 Heidelberg, Germany
[c] Institut für Virologie, Heinrich-Heine-Universität Düsseldorf, D-40225 Düsseldorf, Germany

ARTICLE INFO

Article history:
Received 4 February 2013
Returned to author for revisions
26 February 2013
Accepted 13 March 2013
Available online 3 April 2013

Keywords:
HIV Env C-terminus
Novel Env-CT species
Pyroglutamate modification
Proteolytic processing
Spliced mRNA encoding Env-CT

ABSTRACT

Employing antibodies against the cytoplasmic tail of the HIV-1 glycoprotein (Env-CT), in addition to gp160/gp41, we have identified several novel small Env proteins (< 25 kD) in HIV-1 transfected and infected cells. Mass spectrometric and mutational analyses show that two mechanisms contribute to their generation. Thus the protein, designated Tr-Env-CT (for truncated Env-CT), consists of the C-terminal 139 amino acids (aa) of Env (aa 718–856) with the N-terminal Q718 modified to pyroglutamic acid. It is likely derived from full-length Env protein by proteolytic processing. A further heterogeneous set of slightly larger proteins, termed Env-CT* species, are rather derived from spliced mRNAs containing only those Env C-terminal residues (aa 719–856) which overlap with the second *tat* and *rev* coding exons. They are N-terminally extended in the same reading frame. It is conceivable that essential Env-CT functions may be fulfilled by these novel species rather than by the full-length glycoprotein itself.

© 2013 Elsevier Inc. All rights reserved.

Introduction

A conserved feature of most lentiviral glycoproteins, including the Env glycoprotein of human immunodeficiency virus type 1 (HIV-1), is the presence of a very long cytoplasmic tail region (Env-CT) which, in the case of HIV, is 151 amino acids (aa) long). The existence of an HIV Env-CT region is extremely highly conserved in vivo and, in the simian immunodeficiency virus/monkey model, truncation of the Env-CT leads to suppression of viral replication in vivo (Shacklett et al., 2000). The currently known structural and functional properties of the HIV Env-CT are summarized in two recent reviews (Checkley et al., 2011; Postler and Desrosiers, 2012) and will only be described here briefly. In its central and C-terminal region, the Env-CT contains three regions, termed lentiviral lytic peptides (LLP1-3) regions which exhibit α-helical potential and have been shown to interact with cellular membranes (Chernomordik et al., 1994; Haffar et al., 1991). Presumably mediated by its interaction with the viral matrix protein (MA), the Env-CT has been proposed to play roles in Env incorporation into virions (Freed and Martin, 1996) and in the regulation of fusion between released virions and the membrane of target cells (Murakami et al., 2004; Wyma et al., 2004). Two functional endocytosis motifs are also located in the Env-CT, namely a membrane-proximal tyrosine-based signal (YxxL) and a C-terminal dileucine motif (Env-LL855, 856), both of which operate independently through interaction with clathrin adaptor protein complexes (Berlioz-Torrent et al., 1999; Byland et al., 2007; Wyss et al., 2001). The membrane-proximal YxxL motif has also been shown to play a role in the basolateral targeting of Env (and Gag) in polarised cells (Lodge et al., 1994, 1997). The Env-CT region has also been shown to interact with a number of cellular proteins. Thus, in addition to components of the endocytosis machinery, it has been demonstrated that the Env-CT interacts with calmodulin (Miller et al., 1993), TIP47 (Blot et al., 2003), the guanine nucleotide exchange factor p115-RhoGEF (Zhang et al., 1999), prenylated Rab receptor (Rab1) (Evans et al., 2002), catenin (Kim et al., 1999), luman (Blot et al., 2006) and the dimer of cellular prohibitin 1 and prohibitin 2 (Phb1/Phb2) (Emerson et al., 2010b). In general, the functional importance of these interactions have not been elucidated.

It is important to note that most of the roles presently attributed to the HIV-Env-CT region can be fulfilled by the glycoproteins of other retroviruses, as well by the glycoproteins of enveloped viruses in general, all of which carry much shorter CT regions. This suggests that additional lentiviral Env-CT functions, essential for viral replication and pathogenesis in vivo, still remain to be elucidated.

In this study, we describe several novel Env proteins detected in provirally transfected and infected cells which consist of amino

* Corresponding author. Fax: +49 6221 424932.
E-mail addresses: t.pfeiffer@dkfz.de (T. Pfeiffer), t.ruppert@zmbh.uni-heidelberg.de (T. Ruppert), schaal@uni-duesseldorf.de (H. Schaal), v.bosch@dkfz.de (V. Bosch).

0042-6822/$ - see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.virol.2013.03.009

acid sequences derived from the Env-CT region but lacking the Env membrane-spanning and ectodomains. The existence of these species opens the possibility that essential lentiviral Env-CT specific functions may, in fact, be implemented by these protein moieties rather than by the Env-CT as part of the full-length glycoprotein.

Results

Novel protein moieties containing HIV Env-CT sequences

The HIV Env glycoprotein gp160 is cleaved by the cellular protease furin into the surface gp120 (SU) and the transmembrane gp41 (TM) proteins. The topology of the TM protein predicts that its C-terminal cytoplasmic tail (Env-CT) consists of aa 706–856. A schematic representation of HIV Env illustrating the positions of features of relevance for this study is shown in Fig. 1A. Fig. 1B, lane 1, shows Western blot analysis, employing Chessie 8 mAb (epitope aa 727–732) of a lysate of 293 T cells transfected with wild type proviral plasmid, pNL-Wt. In addition to the expected bands of gp160 and gp41, a number of bands with molecular weights in the range of approximately 13–20 kD were detected. The use of the employed gradient gel system was important for the proper separation of these different Env species. These novel bands were absent employing unrelated mouse monoclonal antibodies, e.g. SIM2 antibodies (against CD4), and were not detected in lysates of untransfected cells. They were also not an artefact of the use of the Chessie 8 mAb but were also detected employing rabbit polyclonal antiserum against gp160 protein, which also contains antibodies against the CT region (data not shown). Further, the same Chessie 8 reactive species could be detected on expression of a different HIV Env protein, namely THRO4156.18 Env (accession number AY835448, (Li et al., 2005)) which exhibits 20% amino acid sequence diversity in comparison to the Env protein employed here (Fig. S1). The construct prec. Env-CT expresses a recombinant protein consisting of the Env-CT region (plus an N-terminal methionine) *alone* (see Fig. 1A). As can be seen in Fig. 1B, lane 9, in line with its calculated size of 17.76 kD, this protein migrated somewhat slower than the 15 kD marker band. Thus the protein in the fastest migrating band in Fig. 1B, lane 1 (running faster than the 15 kD marker) is clearly smaller than the entire Env-CT. It is designated Tr-Env-CT (for truncated Env-CT). The generation of Tr-Env-CT was dependent on expression of full-length Env protein and was not detectable on expression of pNL-ΔEnv with a frame-shift at the beginning of the *env* gene, (Fig. 1B, lane 6), or when Env proteins truncated within the CT region (from proviral pNL-Env-Tr712 (Fig. 1B, lane 3) and pNL-Env-Tr-813 (Fig. 1B, lane 4) or from the subgenomic Env expression plasmid pßAc-Env-Tr752 (Fig. 1B, lane 8)) were expressed. However, the additional larger protein bands, seen on expression of pNL-Wt (lane 1) and migrating slower than the 15 kD marker band, were also detected on expression of pNL-ΔEnv as well as in some of the other lanes in somewhat varying amounts. Since pNL-ΔEnv does not express Env protein at all, these heterogeneous bands (referred to here as Env-CT* bands) were initially regarded as representing some sort of background artefact. However, as will be discussed later, this is not the case but rather the Env-CT* bands turned out to be specific spliced Env-CT isoforms with N-termini which are not derived from Env protein. In Fig. 1C, Western blot analysis of lysates of C8166 T-cells expressing pNL-Wt and pNL-Env-Tr712 after infection with pseudotyped virions is shown. The same Env-CT-related protein species as seen in the transfected 293 T cells were seen, albeit in lower amounts.

Generation of Tr-Env-CT

Our initial focus was on the elucidation of the identity and mode of generation of the Tr-Env-CT protein. Tr-Env-CT was still generated in cells transfected with pNL-Wt incubated in the presence of HIV protease inhibitor, Saquinivir, as well as from the subviral Env expression vector pßAc-Env (Fig. 1B, lanes 2 and 7). The effectiveness of the Saquinivir treatment was confirmed by analysis of HIV Gag protein expression which showed that only precursor $Pr55^{gag}$ and no p24 was present (Fig. S2).

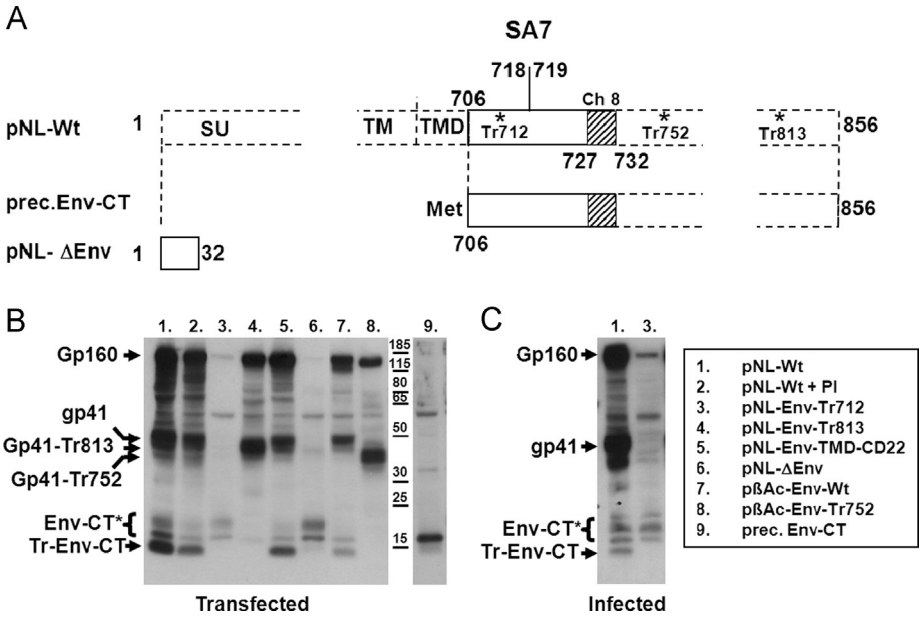

Fig. 1. (A) Schematic representation of the HIV glycoprotein. The region in the Env-CT between Env aa 706 and 732 is expanded. The Chessie 8 epitope between aa 727 and 732 is hatched, the splice acceptor site 7 (SA7) between aa 718 and 719 is marked and the C-termini of the employed Env truncations are shown by stars (*). The dimensions of the Env proteins expressed from prec.Env-CT and pNL-ΔEnv are given below. (B) Western blot analysis of lysates of transfected 293 T cells and of C. C8166 cells infected with VSV-G pseudotyped virions. The plasmids employed in the individual lanes are indicated on the right. In all cases, Env-CT antibodies (Chessie 8) were employed. In B. the positions of molecular weight markers are given and in both (B) and (C) the positions of detected HIV-Env proteins are indicated on the left. Note that in (B) lane 9 is from the same gel as lanes 1–8 but shows a different (shorter) exposure.

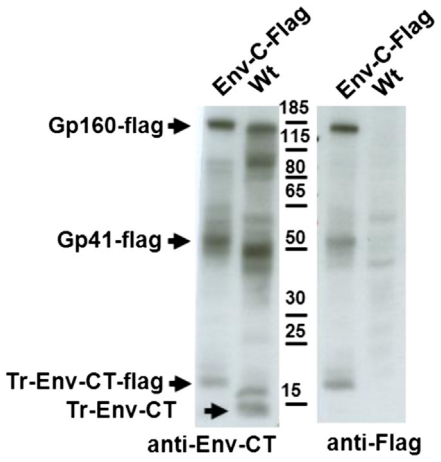

Fig. 2. Western blot analysis of lysates of 293 T cells transfected with pCMV-Env (Wt) and pCMV-Env-C-Flag and probed with Env-CT (Chessie 8) antibodies (left) or anti-Flag antibodies (right). The positions of molecular weight markers are shown in the middle and of detected HIV-Env proteins on the left.

This shows that Tr-Env-CT is not a product of the HIV protease enzyme nor does its generation require HIV sequences outside of pNL4-3 nucleotides 5785–8887. The generation of Tr-Env-CT also does not depend on the sequence of the Env membrane spanning domain (TMD, Env aa 684–705) since it was still produced from pNL-Env-TMD-CD22 in which the HIV-Env-TMD has been replaced with that of the type 1 glycoprotein CD22 (Fig. 1B, lane 5). In order to analyse if Tr-Env-CT contains the Env C-terminus, expression from pCMV-Env-Flag, carrying tandem Flag sequences within 23 additional amino acids at the Env C-terminus, was examined employing Flag antibodies. Flagged gp160 and gp41 (this migrating somewhat slower than unflagged gp41), as well as a band migrating slower than the 15 kD marker, were observed (Fig. 2, right panel). Since this band also additionally reacted with Chessie 8 (Fig. 2, left panel), it represents a flagged version of Tr-Env-CT showing that Tr-Env-CT carries the Env C-terminus.

Determination of the Tr-Env-CT N-terminus

As demonstrated until now, Tr-Env-CT carries both the Env C-terminus as well as the Chessie 8 epitope which is located only 21 aa C-terminal from the Env TMD (see Fig. 1A). Since Tr-Env-CT is smaller than the Env-CT itself (as expressed from prec.Env-CT), its N-terminus is deduced to lie within the Env aa sequence between the TMD and the Chessie 8 epitope (i.e. within aa 706–727). To confirm the identity of Tr-Env-CT and to determine its N-terminus, we used mass spectrometry to analyse the tryptic peptides generated after in-gel digest of purified Tr-Env-CT and gp41 protein bands. To prepare these, Chessie 8 reactive bands were immunoprecipitated from lysates of 2 large (18 cm diameter) dishes of 293 T cells transfected with pßAc-Env-Wt. Initial analysis of small aliquots (2–3% of total) of the precipitated proteins confirmed specific immunoprecipitation with an efficiency of approximately 10–20% of the total (data not shown). The remaining 97–98% of the immunoprecipitate was then applied to a single slot of a gel which was subsequently stained with Coomassie blue. There were no visible bands of gp41 or Tr-Env-CT but, with the orientation provided by the applied marker proteins, gel slices at the respective positions were excised. The tryptic peptides identified in gp41 and Tr-Env-CT, respectively are given in Table 1 and for Tr-Env-CT are shown schematically in Fig. 3B. As to be expected, in the case of the gp41, the peptides were located throughout the protein while most of the identified tryptic peptides in the Tr-Env-CT band were within the Env-CT region.

Table 1
Peptides detected in purified gp41 and Tr-Env-CT proteins.

Peptide	Env aa	Detection	Ratio Tr-Env-CT/gp41
QLLSGIVQQQNNLLR	543–557	Gp41/Tr-Env-CT	0.06
AIEAQQHLLQLTVWGIK*	558–574	Gp41	–
YLKDQQLLGIWGCSGK	586–601	Gp41	–
NEQELLELDK	656–665	Gp41	–
QGYSPLSFQTHLPIPR	710–725	Gp41	–
GPDRPEGIEEEGGER	726–740	Gp41/Tr-Env-CT	1.4
LVNGSLALIWDDLR*	748–761	Gp41/Tr-Env-CT	1.1
SLCLFSYHR	762–770	Gp41/Tr-Env-CT	1.7
LRDLLLIVTR*	771–780	Gp41/Tr-Env-CT	1.5
DLLLIVTR	773–780	Gp41/Tr-Env-CT	1.3
RGWEALK	788–795	Tr-Env-CT	–
NSAVSLLNATAIAVAEGTDR*	809–828	Gp41/Tr-Env-CT	0.6
VIEVVQGAYR	829–838	Gp41/Tr-Env-CT	0.8

The detected peptides (sequenced by MS/MS), their positions within the Env protein sequence and whether detected in purified gp41, Tr-Env-CT or both is given. Peptide intensities were calculated by the label free quantitation tool of the MaxQuant Software. In the right column the ratios of the intensities of the individual peptides from gp41 in comparison to the intensities of the same peptides from Tr-Env-CT are given. Peptides marked with an asterisk were detected in different charge states and the average Tr-Env-CT/gp41 ratio taken.

An exception was a single peptide from the gp41 ectodomain which was presumably a contaminant from a breakdown product of gp41 or gp160 present in the same gel slot. The ratio of the individual peptide intensities derived from the gp41 and Tr-Env-CT proteins (Table 1) supports this conclusion. The value for the contaminating gp41 ectodomain peptide was much lower (0.06) than the values for all of the other shared peptides (0.6–1.7). These latter values further indicate that, at any rate in this transfected 293 T cell lysate, the molar amount of Tr-Env-CT was in the same range as that of gp41.

The most N-terminal tryptic peptide identified in Tr-Env-CT was located at aa 726–740 (see Table 1 and Fig. 3B) whereas in the case of gp41 further N-terminal peptides including that at position 710–725 (QGYSPLSFQTHLPIPR) could be identified (Table 1). We therefore subsequently concentrated on the identification of a Tr-Env-CT peptide starting with a non-tryptic N-terminus and containing aa sequences N-terminal to Env aa 726. Using the error tolerant search option of the Mascot software, which tolerates a single unknown modification within a peptide, the peptide QTHLPIPR could be identified with high Mascot score. This identification indicated that the N-terminus of Tr-Env-CT is located at Env residue Q718. This means that Tr-Env-CT is 139 aa long (see Fig. 3B) and is thus 13 aa (12 aa plus initiation methionine) shorter than the Env-CT protein expressed from prec.Env-CT. This is consistant with its gel mobility. The peptide mass of the QTHLPIPR peptide and the series of y-fragments (the peptide fragments generated during sequencing) of this peptide (shown in Fig. 3A) indicated, however, that the N-terminal glutamine residue had been modified to pyroglutamic acid (pGlu). This modification is catalysed by cellular glutaminyl cyclase (QC), of which two different isoenzymes exist. In the cells of relevance here (transfected 293 T kidney cells and infected T-lymphocytes) the Golgi-localised isoQC enzyme is likely to be responsible. In general, modification of N-terminal glutamine to pGlu has been described to increase protein (or peptide) stability and biological activity.

A likely hypothesis to account for the generation of Tr-Env-CT is that it is derived from full-length Env protein by proteolytic processing. Fig. 3C illustrates the aa sequence in the vicinity of the N-terminal pGlu of Tr-Env-CT. The percentage conservation of each residue (as evaluated using the Quickalign tool with reference to the 2011 web alignment provided by the HIV Database, Los Alamos, New Mexico, USA (http://www.hiv.lanl.gov/) is shown and illustrates that a long stretch of highly conserved aa is located

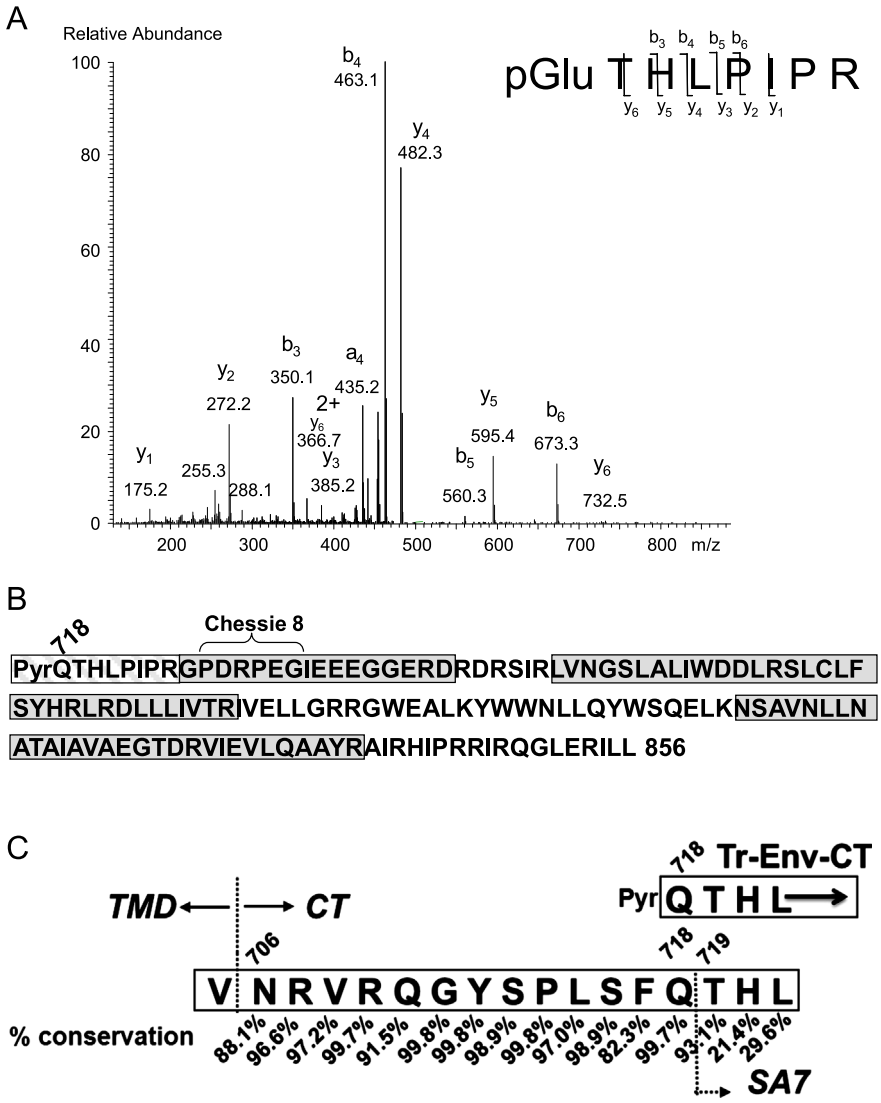

Fig. 3. (A) MSMS spectrum of 718-pGluTHLPIPR-725. A doubly protonated peptide with $m/z=472.77$ was selected for fragmentation. The sequence …HLPIPR is clearly detected by the y-fragments 1 to 6. The difference between the peptide mass and the mass of the y6 fragment can only be explained by an N-terminal glutamine residue that had been modified to pyroglutamic acid (pGlu) given that the second amino acid threonine is correct. (B) Derived amino acid sequence of Tr-Env-CT (Env aa 718-856). The grey blocks indicate regions covered by identified tryptic peptides and the hatched grey box the N-terminal peptide with modified glutamine 718 (pyr-Q). The amino acids in the Chessie 8 epitope are indicated. (C) Amino acid sequence and percentage conservation of the Env aa in the vicinity of Env Q718 as determined using the Quickalign tool with reference to the 2011 web alignment provided by the HIV Database, Los Alamos, New Mexico, USA (http://www.hiv.lanl.gov/). The border of the Env membrane anchor (TMD) and the cytoplasmic domain (CT) as well as the junction of splice acceptor (SA7) are shown. The position of the N-terminus of the Tr-Env-CT protein is indicated.

N-terminal of pGlu. However, consensus cleavage motifs for known processing proteases (searched using the tool http://web.expasy.org/peptide_cutter) could not be identified within this region and mutation of individual aa did not prevent generation of Tr-Env-CT (not shown). Nevertheless, since numerous cellular proteases lack consensus cleavage motifs but rather recognise structural motifs, we still favour proteolytic processing as the most likely mechanism for the generation of Tr-Env-CT. Future studies may throw more light on these issues.

Env-CT amounts increased in the absence of rev.*

As documented in Fig. 1B, the slightly larger set of heterologous species, designated Env-CT*, was also generated from pNL-ΔEnv which, as a result of a frame-shift mutation at the beginning of the *env* gene, fails to express Env proteins (and also does not express Tr-Env-CT). It was thus conceivable that the Env-CT* species were expressed from spliced mRNAs. In order to investigate this, expression from pNL-(rev-), encoding a defective truncated Rev protein (Rev-Tr62) was analysed. In the absence of functional Rev, only spliced HIV mRNAs are efficiently transported to the cytoplasm and translated (Dayton et al., 1988; Malim et al., 1988). Importantly, in pNL-(rev-), the *env* gene is intact and, in principle, encodes only minimally altered Env protein. As expected, pNL-(rev-) expresses neither Gag nor gp160 and gp41 proteins (Fig. 4A). However, Env-CT* species were strongly increased pointing to their production being by a mechanism independent of Env expression, most probably by expression from fully spliced Rev-independent mRNA.

Expression of Env-CT species from prespliced (SD4–SD7) vector*

The Env aa sequence 719–856, containing the Chessie 8 epitope, is located downstream of splice acceptor 7 (see Fig. 1A), and is present on nearly all HIV mRNA species whether full-length, partially or fully spliced. In the majority of fully spliced mRNAs (encoding Tat and Rev), splice donor 4 (SD4) has been joined to splice acceptor 7 (SA7). pLTRctat/crev, schematically depicted in

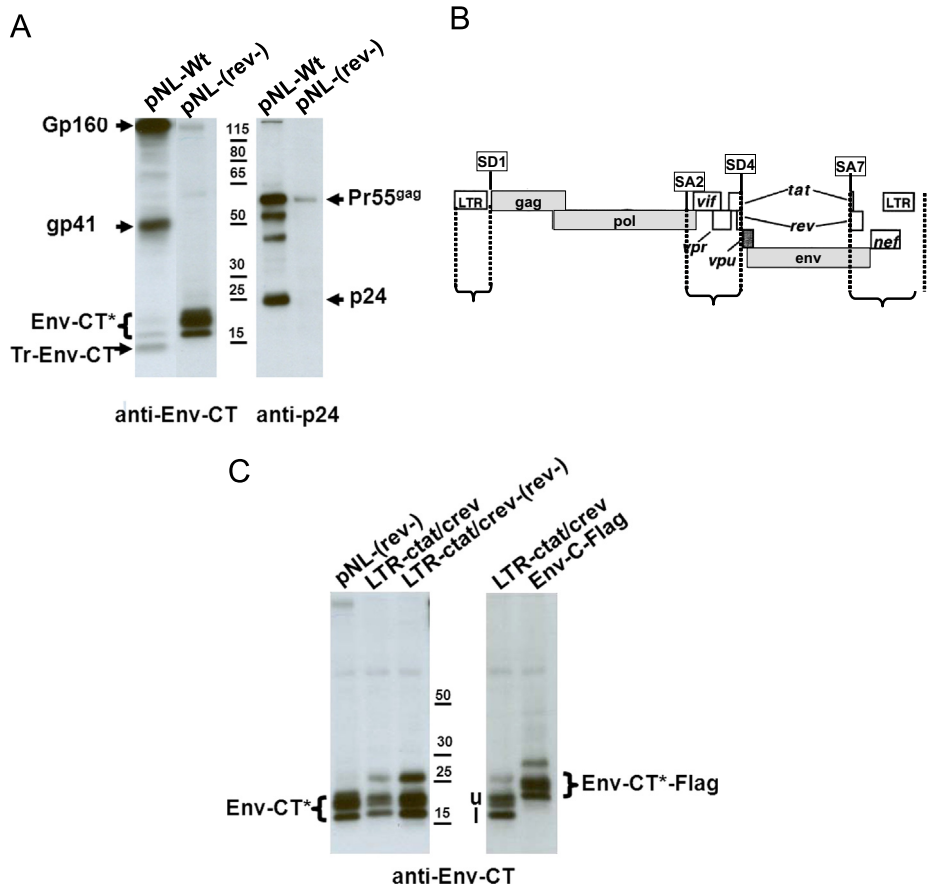

Fig. 4. (A) Western blot analysis of lysates of 293 T cells transfected with pNL-Wt and the rev-defective derivative pNL-(rev-). The antibodies employed are indicated below the respective images. (B) Schematic depiction of the construct pLTR-ctat/crev. The entire HIV genome is given with the positions of relevant splice donor (SD) and splice acceptor (SA) sites indicated. The HIV sequences remaining in pLTR-ctat/crev are shown in brackets below. (C) Western blot analysis with Chessie 8 antibodies of 293 T cells transfected with pNL-Wt, pLTR-ctat/crev and the derivative pLTR-ctat/crev-(rev-) (left blot) and with pLTR-ctat/crev and the flagged derivative pLTR-ctat/crev-Flag (right blot). The positions of the upper (u) and lower (l) regions of the Env-CT* species which have been separately analysed by mass spectrometry are indicated.

Fig. 4B, is a subgenomic construct mimicking this SD4–SA7 splicing event. In this construct, splice donor 1 (SD1) has been recombinantly joined to splice acceptor 2 (SA2) and SD4 to SA7 (Schaal et al., 1993). The construct encodes Vif and Vpr and, as a result of recombinant splicing between SD4 and SA7, expresses the spliced gene products Tat and Rev. The derivative pLTRctat/crev-(rev-) encodes defective truncated Rev protein (Rev-Tr62). Western blot analysis of lysates of cells transfected with these constructs employing Chessie 8 antibodies is shown Fig. 4C, left panel. A similar set of Env-CT* species as seen in a lysate of cells transfected with pNL-(rev-) (plus an additional slightly larger band present in variable amounts) were detected and again the expression of these protein entities was independent of Rev expression (Fig. 4C, left panel). In order to study if these protein entities contained the Env C-terminus, lysates of cells transfected with pLTRctat/crev-Flag carrying 23 additional C-terminal aa (with tandem Flag epitopes) were examined. Expressed proteins reacted with both Flag antibodies (not shown) and Chessie 8 antibodies and migrated slightly slower than the unflagged versions (Fig. 4C, right panel), which shows that they all contain the HIV-Env C-terminus (Fig. 4C, right panel).

Narrowing down the identity of the Env-CT species*

The Env-CT* species contain the Env-C-terminus and the Chessie 8 epitope and are larger (migrate more slowly in gel) than Tr-Env-CT. This means that the Env-CT* species are expected to contain additional N-terminal amino acids upstream of Env aa 719–856 which are encoded in pLTR-ctat/crev. Thus, in order to gain further information as to their identities, mass spectrometric analysis of purified Env-CT* species was performed. Immunoprecipitation from lysates of pLTRctat/crev transfected cells was carried out as previously described for gp41 and Tr-Env-CT. Two gel slices containing the upper (u) and lower (l) Env-CT* species (as labelled in Fig. 4C) were excised, trypsin digestion performed and the aa sequences of the identified peptides screened against all the HIV open reading frames (ORF) present in pLTRctat/crev. In Fig. 5A, a region of the ORF potentially encoding the Env-CT sequences is shown. The region between a stop codon at position +45 (i.e. upstream from Env aa 719) and the Env stop codon is depicted. Presumably due to limitations in amounts of protein, only a few peptides could be identified and only in the upper Env-CT* band. However, all of the identified peptides were in the depicted ORF. Three of the peptides were entirely within the Env-CT but in the case of the identified tryptic peptide ATHLPIPR, the N-terminal alanine residue, is upstream (+1 position) of the Env-CT sequence (and is preceded by a K residue). The fragment spectrum of this peptide is shown in Fig. 5B. This shows that amino acids in the same ORF but upstream of Env aa 739 are contributing to the larger size of the Env-CT* species. In order to confirm this involvement of upstream in-frame aa, we introduced a stop codon in the upstream region, namely at K+2. Fig. 5C shows a Western blot of lysates of cells transfected with pLTR-ctat/crev and this derivative. The introduced stop codon

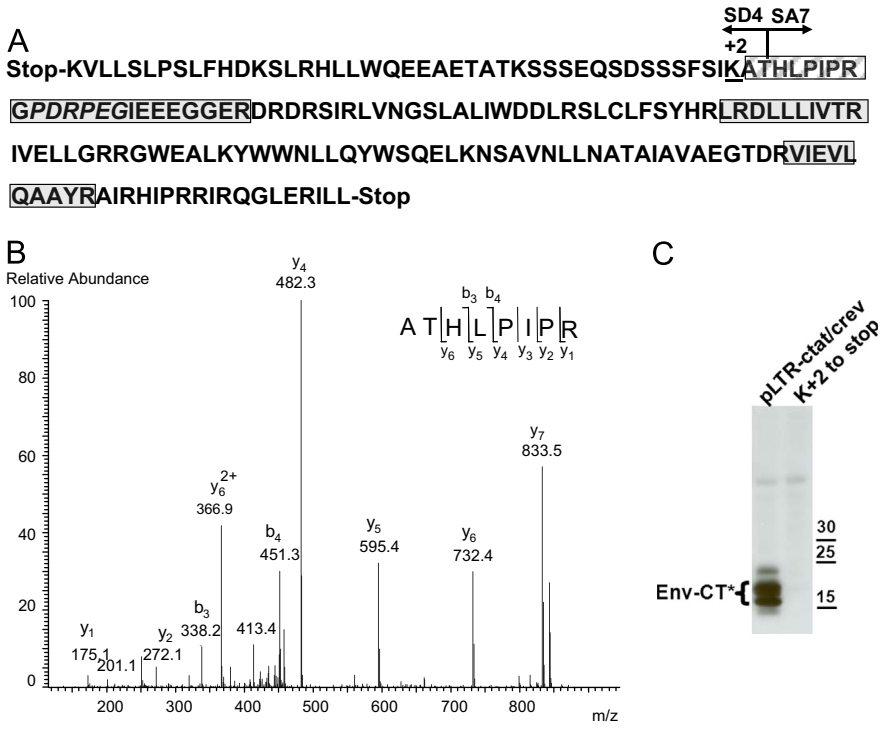

Fig. 5. (A) Amino acid sequence of the ORF in pLTR-ctat/crev encoding 138aa of the Env-CT. The sequence between the first upstream stop codon and the Env stop codon is depicted. The junction between SD4 and SA7 is shown. The Env-CT tryptic peptides identified by MSMS sequencing are boxed in grey. The peptide ATHLPIPR which spans the SD4-SA7 junction is boxed in hatched grey. The mutated aa in pLTR-ctat/crev (K+2-stop) is underlined and labelled. (B) MSMS spectrum of peptide ATHLPIPR. A doubly protonated peptide with m/z=452.77 was selected for fragmentation. The sequence …HLPIPR is clearly detected by the y-fragments 1 to 6. The difference between peptide mass and the mass of the y6 fragment can only be explained by an N-terminal alanine given that the second amino acid threonine is correct. (C) Western blot analysis, employing Env-CT (Chessie 8) antibodies of pLTR-ctat/crev and the derivative pLTR-ctat/crev-(Env-CT+2 stop). The position of the Env-CT* proteins is given on the left and of molecular weight markers in the middle.

at the +2 position led to abrogation of expression of all of the Env-CT* moieties confirming that they contain N-terminal aa in the upstream reading frame.

Discussion

In this study we describe the detection and preliminary analysis of several novel protein entities containing amino acid sequences from the HIV Env cytoplasmic C-terminus. The product Tr-Env-CT has been demonstrated to consist of the last 139 aa of Env (Env aa 718–856) in which the N-terminal aa Q718 has been modified to pyroglutamic acid (5′-oxoproline (pGlu)) (see Fig. 6). N-terminal pGlu modification is present on a number of peptide hormones, notably on thyrotropin-releasing hormone (TRH) and gonadotropin-releasing hormone (GnRH), as well as on secretory proteins such as neurotensin and fibronectin (Vale et al., 1981). Furthermore, significant amounts of beta-amyloid (Aß) peptides, involved in Alzheimer's disease, carry N-terminal pGlu (Saido et al., 1995, 1996). The tissue distribution of the enzyme catalysing this modification, glutaminyl cyclase (QC), was initially described to be chiefly restricted to regions of the brain (Pohl et al., 1991). However, subsequently, an isoenzyme (isoQC) with similar activity but only about 45% sequence homology was described to have virtually ubiquitous tissue distribution (Cynis et al., 2008). The isoQC enzyme is subcellularly localised to the Golgi complex where it is believed to play a general role in the protein-maturation machinery of the secretory pathway (Cynis et al., 2008). The general consequence of N-terminal modification to pGlu is to increase peptide or protein stability and increase biological activity. In this study, only N-terminally modified (and no unmodified) QTHLPIPR peptide was detected which strongly points to the modification being catalysed intracellularly by isoQC. At present, no biological activity has been assigned to the Tr-Env-CT protein. However, if this becomes available, the use of inhibitors of QC enzyme may be employed to analyse the possible involvement of the N-terminal pGlu.

As shown here, expression of full-length Env glycoprotein is a prerequisite for the generation of Tr-Env-CT making it likely that it is a product of proteolytic processing of the full-length molecule. However the HIV protease itself does not appear to be the processing molecule. This is in contrast to the situations with the lentivirus equine infectious anemia virus (EIAV) and the further retroviruses murine leukemia virus (MuLV) and Mason Pfizer virus (MPV). In these cases, proteolytic removal of the C-terminal domains of the respective transmembranal glycoproteins is carried out by the respective viral protease enzymes and is required to activate Env-mediated membrane fusion in virus particles (Brody et al., 1994; Green et al., 1981; Rein et al., 1994; Rice et al., 1990). Furthermore, in the case of certain mutant HIVs, selected for resistance to the cholesterol-binding compound amphotericin B methyl ester (AME), the viral protease enzyme does actually cleave the HIV-Env-CT region from the transmembranal gp41 protein (Waheed et al., 2007). However in the case of Tr-Env-CT entity described here, since the HIV protease does not appear to play a role, it is likely that cleavage is carried out by a cellular protease. Since modification of Q718 by isoQC would likely occur within the Golgi complex, proteolytic cleavage to (initially unmodified) Tr-Env-CT would be assumed either to take place there or prior to Env transport from the endoplasmatic reticulum. Proteolytic cleavage of gp160 to gp120 and gp41 also takes place within the Golgi complex making it possible that Tr-Env-CT is derived from either gp160, gp41 or both. We have shown that Tr-Env-CT is still generated when non-cleavable gp160, with a

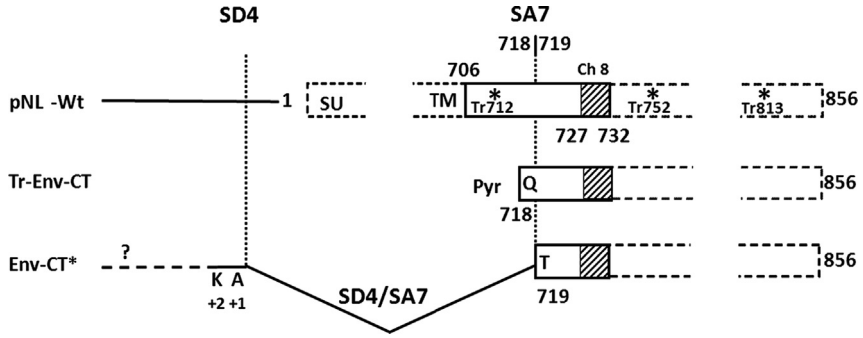

Fig. 6. Schematic representation of the Env-CT species described in this paper. In the HIV glycoprotein, the region in the Env-CT between Env aa 706 and 732 is expanded. The Chessie 8 epitope between aa 727 and 732 is hatched and the position of the splice acceptor site 7 (SA7) between aa 718 and 719 is shown. Splicing between splice donor site 4 (SD4, nucleotides 6046/6047) and SA7 is schematically depicted. The dimensions of Tr-Env-CT and the spliced Env-CT* species are shown schematically. The positions of A+1, detected by mass spectrometry, and K+2, which has been mutated to stop, are shown. The unknown translation start position of the Env-CT* species is indicated by a question marks (?).

mutation at the processing site between gp120 and gp41 (Bosch and Pawlita, 1990) was expressed (data not shown) but this does not necessarily need to hold true in the wild type situation.

Concerning the amino acid sequence requirement for Tr-Env-CT generation, we have demonstrated that the sequence of the HIV Env membrane spanning domain, which is located only 12 aa N-terminal to Q718, is not of importance. The 12 amino acids upstream of Q718 and prior to the Env membrane anchor are very highly conserved in all HIV-1 Envs and it is conceivable that they form part of a protease recognition domain. We have mutated these amino acids and further amino acids surrounding Q718 itself and analysed the properties of the mutated Envs. These studies were, however, hampered by the fact that the sequence for splicing at splice acceptor 7 (SA7, to the second *tat* and *rev* coding exons) overlaps with Q718 and several other neighbouring codons. Actually, the analyses did not yield further information since the introduced changes either did not prevent generation of Tr-Env-CT or adversely influenced splicing at SA7 (data not shown). However, it has to be remembered that many proteases actually do not have strict substrate sequence requirements but rather recognize conformational features. Finally though, it cannot strictly be ruled out that the Tr-Env-CT entity may be the product of unspecific proteolysis. However, as shown in Table 1, evaluation of the relative intensities of peptides isolated from both the purified gp41 and Tr-Env-CT proteins showed that, at any rate in transfected 293 T cells, the amount of Tr-Env-CT is not much lower than that of gp41. This rather indicates, but does not yet prove, that Tr-Env-CT is a specific product and presumably serves a presently unknown function during HIV infection in vivo.

Since the slightly larger species, designated Env-CT*, are independent (Fig. 1) of expression of Env protein, it was likely that they were products of spliced HIV mRNAs. Confirming this, we observed that proviral pNL-(rev-), encoding mutated, non-functional Rev protein and failing to express Gag and full-length Env proteins, expressed considerably increased amounts of the heterogeneous Env-CT* species. This suggested that Env-CT* species may be expressed from spliced mRNAs which are predominantly transported to the cytoplasm in the absence of Rev. Confirming this, Env-CT* species were also expressed from the subgenomic construct pLTRctat/crev in which SD4 has been fused to SA7 (i.e. to the second coding exons of *tat* and *rev*). In this construct, Env sequences from aa 1–718 are absent but, as is the case for all natural *tat* and *rev* transcripts, the last 138 C-terminal aa of Env (719–856), containing the Chessie 8 epitope, are present and overlap with the *tat* and *rev* second coding exons.

The questions which were then of interest concerned first, the identities of the different Env-CT* species and second, how they are expressed. We have performed preliminary investigation of these issues. As a first step, we could show that all of the Env-CT* species contain the Env protein C-terminus since they were all increased in size when a Flag sequence was fused to the Env C-terminus. Since all of the Env-CT* species are larger than Tr-Env-CT (Env aa 718–856), it is to be expected that they contain additional non-Env N-terminal aa. In the tryptic peptide, ATHLPIPR, which was identified from the upper Env-CT* band by mass spectrometry, the N-terminal alanine is derived from the in-frame sequence upstream of the Env sequence. The presence of N-terminal in-frame amino acids upstream of Env aa T719 is presumably to be expected since in pLTR-ctat/crev SD4 is fused to SA7. Due to limitations in protein amounts, no further information could be derived from these mass spectrometric analyses. However, by introducing a stop codon at the +2 position in the N-terminal in-frame region (K+2), we could confirm that the Env-CT* species carry N-terminal in-frame aa. In Fig. 6, the position of the residues in Env-CT* within the HIV genome are schematically shown. Concerning translational initiation of Env-CT*, it is of note that there are no potential initiation methionines in the in-frame upstream region (see Fig. 5A). Thus it is likely that the Env-CT* species are expressed from one or several multiply spliced mRNAs i.e. from transcripts which have been additionally spliced upstream from SD4. Expression from *different* multiply spliced transcripts could explain the fact that Env-CT* consists of distinct species. On the other hand, it cannot be ruled out that post-translational processing or modification of initially translated species could account, at least in part, for the heterogeneity observed. These issues are beyond the scope of the present study but remain to be addressed in the future.

At steady state, Env-CT* species expression from wild type provirus is not very high (Fig. 1B) so that it cannot be ruled out that its detection from the mutated and subgenomic plasmids described here is only an artefact. On the other hand, the expression of the set of Env-CT* proteins is very robust in the absence of Rev protein (Fig. 4). This raises the possibility that at early time points post-infection, before Rev levels are high, Env-CT* species may be expressed and be of importance.

At present we do not have any information as to the possible roles, if any, which the novel Env-CT species described here may play in HIV infection in vivo. As mentioned in the Introduction, several cellular proteins interact with the Env-CT domain and it is possible that these proteins functionally interact with Tr-Env-CT or Env-CT* species rather than with full-length Env protein. We have previously shown that the heterodimer of prohibitin 1 and prohibitin 2 (Phb1/Phb2) is an interaction partner of the Env-CT and, in fact, in coimmunoprecipitation experiments it bound more

strongly to expressed Env-CT alone (with a membrane anchor) than to full-length Env (Emerson et al., 2010b). Future experiments are required to shed further light on these issues.

Methods

Constructs

We have employed the proviral construct pNL4-3^{BH10env} (Wilk et al., 1992), referred to here as pNL-Wt, in this study. Nucleotide numbering given here is from pNL4-3 (Adachi et al., 1986) and Env amino acid (aa) numbering from strain BH10-Env (Ratner et al., 1985). Due to introduced stop codons, pNL-Env-Tr712 and pNL-Env-Tr813 encode for truncated Env proteins of 712 and 813 aa, respectively (Wilk et al., 1992). In pNL-TMD-CD22, the HIV-Env transmembrane domain (aa 684–705) has been replaced with that of the type 1 membrane protein CD22 (Wilk et al., 1996). pNL-ΔEnv does not express Env protein due to an introduced frame-shift at the beginning of the *env* gene (Henriksson and Bosch, 1998). pNL-(rev-) encodes defective truncated Rev protein (Rev-Tr62). It was generated employing the oligonucleotide 5′ CGGATCCT**C**AGC**TG**TTA**G**CTGGGACG 3′ (mismatched nucleotides underlined in bold,) which also introduces a PvuII restriction site and 3 conservative aa changes within the overlapping Env protein. The provirally derived construct, referred to here as pLTRctat/crev, expresses Tat and Rev proteins from cDNAs and has been described previously (Schaal et al., 1993). In this construct splice donor 1 (SD1) has been joined to splice acceptor 2 (SA2) and SD4 to SA7. The derivative pLTR-ctat/crev (rev-) encodes Rev protein with the same mutation as described above. In the derivative pLTR-ctat/crev-Flag, three tandem Flag epitopes (total aa sequence DYKDHDGDYKDHDIDYKDDDDKA) have been introduced in frame at the C-terminus of the sequence potentially encoding the C-terminal 138 aa of the Env-CT (Env aa 719–856). In the further derivative pLTR-ctat/crev-(Env-CT+2 stop), the potential K residue two aa upstream (i.e. 5′) of the Env-CT sequence has been mutated to stop (TAG) (see Fig. 5A). pßAc-Env-Wt is a subviral expression vector for strain BH10-Env, (Sal1-Xho1 fragment from pNL4-3^{BH10env}, nucleotides 5785–8887 (pNL4-3 numbering)) under control of the ß-actin promoter (Krausslich et al., 1993) while the derivative pßAc-Env-Tr752 encodes truncated Env protein of 752 aa (Wilk et al., 1992). A further subviral Env expression vector is pCMV-Env, again containing the HIV Sal1-Xho1 fragment, in this case under control of the CMV immediate early promoter. The derivative pCMV-Env-Flag has been generated by introducing three tandem Flag epitopes in frame at the C-terminus of the Env protein. Env expression from pCMV-Env is enhanced by coexpression of Rev protein which was achieved from pCMV-srev (Mermer et al., 1990). pTHRO4156 clone 18 (accession number AY835448) (Li et al., 2005) is an expression vector for subtype B Env glycoprotein with 20% diversity as compared to BH10 Env. It was obtained from the NIH AIDS Research and Reference Reagent programme. The construct prec.Env-CT was generated by standard PCR procedures and encodes an N-terminal start methionine followed by Env aa 706–856 under control of the EF1-α promoter in the vector pWPXLd (kindly provided by D. Trono, Geneva, Switzerland) from which the gene for GFP had been excised. pMD.G is an expression vector for the G glycoprotein of vesicular stomatitis virus (VSV-G) (Naldini et al., 1996).

Protein expression

293 T cells were cultivated in Dulbecco's modified Eagle's medium, 10% fetal calf serum (FCS) and C8166 T-cells in RPMI medium, 10% FCS. 293 T cells were transfected with the respective constructs, in one experiment in the presence of 3 μM HIV protease inhibitor Saquinavir, and at 48 h post-transfection (p.t.), cell lysates were prepared in gel electrophoresis sample buffer (containing 2% SDS, 5% mercaptoethanol). In order to efficiently express pNL-Env-Tr712 (and pNL-Wt) in H9 cells, pseudotyped Wt and mutant virions were generated in 293 T cells by cotransfection of the respective proviral constructs with pMD.G. Forty-eight hours p.t., aliquots of culture supernatants were used to infect H9 T-cells and after removal of input virions and washing, the infected cells were further incubated for 48 h and, subsequently cell lysates prepared. Aliquots of cell lysates were separated on precast NuPAGE Bis–Tris 4–12% polyacrylamide gels run with the supplied MES SDS running buffer (Novex Life Technologies). Western blot analysis employing mouse anti-gp41 (Chessie 8) recognising an epitope within the Env-CT (aa 727–732 (PDRPEG) of LAI gp160 (Abacioglu et al., 1994)), rabbit anti-gp160 serum prepared in our laboratory against gel purified gp160 protein expressed in insect cells, anti-p24 (183-H12-5C), anti-Flag monoclonal antibody M2 (Sigma-Aldrich, USA) or mouse anti-CD4 mAb SIM2 was performed as previously described (Emerson et al., 2010a).

Immunoprecipitation of proteins for mass spectrometric analysis

Two large plates (18 cm diameter) of 293 T cells were transfected with either pßAc-Env-Wt or pLTRctat/crev. Forty-eight hours p.t., the cells were suspended in medium (this achieved by strong pipetting), centrifuged, washed once with 10 ml phosphate buffered saline (PBS) and the cell pellets taken up in 1% Triton-X-100 in PBS containing protease inhibitor cocktail (Roche Diagnostics, Mannheim, Germany). After incubation for 10 min. on ice, the lysates were clarified by centrifugation at 2700 × *g* for 5 min. and the supernatants used further.

Prior to generation of the lysates, aliquots of washed Gamma-Bind Plus Sepharose beads (GE Healthcare Biosciences, Uppsala, Sweden) (each equivalent to 100 μl slurry) were incubated with either 500 μl SIM2 (containing mAbs against CD4) or anti-gp41 Chessie 8 hybridoma supernatants (each containing approx. 5 μg IgG) plus 500 μl 1% Triton-X-100 in PBS for 5 h at 4 °C with gentle agitation and subsequently washed. To reduce potential unspecific binding, cell lysates were first incubated for 1 h with SIM2-bound beads and, after removal of these by centrifugation, overnight at 4 °C with Chessie 8-bound beads. Both the SIM2- and the Chessie 8-beads were washed five times with 1% Triton-X-100 in PBS and bound proteins dissolved in gel electrophoresis sample buffer. Aliquots of the original lysates and of the unspecifically (SIM2) and specifically (Chessie 8) bound proteins were subjected to Western blot analysis to confirm and quantify specific immunoprecipitation (in general 10–20% of the Chessie 8 reactive proteins in the lysates were specifically immunoprecipitated). The total remaining amount (>95%) of the proteins bound to the Chessie 8 beads were then applied to a gel and stained with Coomassie blue. With guidance from applied marker proteins, slices of the gel at the positions of the proteins of interest were excised for further use.

Mass spectrometric analysis

Excised gel pieces were transferred to a 96-well plate and reduced, alkylated and digested with trypsin as described previously (Catrein et al., 2005). Peptides were extracted from the gel pieces with 50% acetonitrile/0.1% TFA, concentrated almost to dryness in a vacuum centrifuge and diluted to a total volume of 30μl with 0.1% TFA. 25 μl of the sample were analysed in a nanoHPLC system (Eksigent Technologies, Axel Semrau GmbH & Co., 45549 Sprockhövel, Germany) coupled to an ESI LTQ Orbitrap mass spectrometer (ThermoFisher Scientific). The sample was

loaded on a C18 trapping column (Inertsil, LC Packings, GL Sciences) with a flow rate of 10 μl 0.1% TFA/min.. Peptides were eluted and separated on an analytical column (75 μm × 150 mm) packed with Inertsil 3 μm C18 material (LC Packings, GL Sciences) with a flow rate of 200 nl/min in a gradient of buffer A (0.1% formic acid) and buffer B (0.1% formic acid in acetonitrile): 0–6 min: 0–3% buffer B; 6–60 min: 3–40% buffer B; 60–65 min: 60–90% buffer B. One survey scan (res: 60000) was followed by 5 information dependent product ion scans. Peptide sequencing by MS/MS was performed in the ion trap analyser. Two+(2+), 3+ and 4+ charged ions were selected for fragmentation.

Tandem mass spectrometry (MS/MS) spectra were searched against a small protein database (representing approximately 300 random database entries) containing, in the case of Tr-Env-CT, the aa sequence of the Env protein of HIV strain BH10 (Ratner et al., 1985) and, in the case of Env-CT*, the aa sequences of all the HIV open reading frames (ORF) present in pLTR-ctat/crev (Schaal et al., 1993) using the Mascot software (Matrix Science). The algorithm was set to use semitrypsin as proteolytic specificity, which means that only one trypsin cleavage site was required for peptide generation. Carbamidomethyl was set as a fixed modification of cysteine, and oxidized methionine and deamidation of asparagines and glutamines as variable modifications. Mass tolerance was set to 5 ppm and 0.5 Da for MS and MS/MS, respectively. After a first database search the error tolerant option of the Mascot software was used to identify peptides carrying unknown modifications. Peptides with a score value below a 5% probability to be false positive were taken as good candidates and the MS/MS spectra of these peptides were check manually. For label free peptide quantitation the MaxQuant software (version 1.3.0.5) (Cox and Mann, 2008) was used with the default parameters except that the peptide false discovery rate was set to 5%, and the minimal ion score was set to 30.

Acknowledgments

We thank Martin Löchelt and Oliver Fackler, Heidelberg and Marek Widera, Düsseldorf for fruitful discussion. Monoclonal antibodies against gp41 (Chessie 8) and against p24 (H183-H12-5C) were obtained from the NIH AIDS Research and Reference Reagent Program.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at http://dx.doi.org/10.1016/j.virol.2013.03.009.

References

Abacioglu, Y.H., Fouts, T.R., Laman, J.D., Claassen, E., Pincus, S.H., Moore, J.P., Roby, C. A., Kamin-Lewis, R., Lewis, G.K., 1994. Epitope mapping and topology of baculovirus-expressed HIV-1 gp160 determined with a panel of murine monoclonal antibodies. AIDS Res. Hum. Retroviruses 10, 371–381.

Adachi, A., Gendelman, H.E., Koenig, S., Folks, T., Willey, R., Rabson, A., Martin, M.A., 1986. Production of acquired immunodeficiency syndrome-associated retrovirus in human and nonhuman cells transfected with an infectious molecular clone. J. Virol. 59, 284–291.

Berlioz-Torrent, C., Shacklett, B.L., Erdtmann, L., Delamarre, L., Bouchaert, I., Sonigo, P., Dokhelar, M.C., Benarous, R., 1999. Interactions of the cytoplasmic domains of human and simian retroviral transmembrane proteins with components of the clathrin adaptor complexes modulate intracellular and cell surface expression of envelope glycoproteins. J. Virol. 73, 1350–1361.

Blot, G., Janvier, K., Le Panse, S., Benarous, R., Berlioz-Torrent, C., 2003. Targeting of the human immunodeficiency virus type 1 envelope to the trans-Golgi network through binding to TIP47 is required for env incorporation into virions and infectivity. J. Virol. 77, 6931–6945.

Blot, G., Lopez-Verges, S., Treand, C., Kubat, N.J., Delcroix-Genete, D., Emiliani, S., Benarous, R., Berlioz-Torrent, C., 2006. Luman, a new partner of HIV-1 TMgp41, interferes with Tat-mediated transcription of the HIV-1 LTR. J. Mol. Biol. 364, 1034–1047.

Bosch, V., Pawlita, M., 1990. Mutational analysis of the human immunodeficiency virus type 1 env gene product proteolytic cleavage site. J. Virol. 64, 2337–2344.

Brody, B.A., Rhee, S.S., Hunter, E., 1994. Postassembly cleavage of a retroviral glycoprotein cytoplasmic domain removes a necessary incorporation signal and activates fusion activity. J. Virol. 68, 4620–4627.

Byland, R., Vance, P.J., Hoxie, J.A., Marsh, M., 2007. A conserved dileucine motif mediates clathrin and AP-2-dependent endocytosis of the HIV-1 envelope protein. Mol. Biol. Cell 18, 414–425.

Catrein, I., Herrmann, R., Bosserhoff, A., Ruppert, T., 2005. Experimental proof for a signal peptidase I like activity in Mycoplasma pneumoniae, but absence of a gene encoding a conserved bacterial type I SPase. FEBS J. 272, 2892–2900.

Checkley, M.A., Luttge, B.G., Freed, E.O., 2011. HIV-1 envelope glycoprotein biosynthesis, trafficking, and incorporation. J. Mol. Biol. 410, 582–608.

Chernomordik, L., Chanturiya, A.N., Suss-Toby, E., Nora, E., Zimmerberg, J., 1994. An amphipathic peptide from the C-terminal region of the human immunodeficiency virus envelope glycoprotein causes pore formation in membranes J. Virol. 68, 7115–7123.

Cox, J., Mann, M., 2008. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. Nat. Biotechnol. 26, 1367–1372.

Cynis, H., Rahfeld, J.U., Stephan, A., Kehlen, A., Koch, B., Wermann, M., Demuth, H.U., Schilling, S., 2008. Isolation of an isoenzyme of human glutaminyl cyclase: retention in the Golgi complex suggests involvement in the protein maturation machinery. J. Mol. Biol. 379, 966–980.

Dayton, A.I., Terwilliger, E.F., Potz, J., Kowalski, M., Sodroski, J.G., Haseltine, W.A., 1988. Cis-acting sequences responsive to the rev gene product of the human immunodeficiency virus. J. Acquir. Immune Defic. Syndr. 1, 441–452.

Emerson, V., Haller, C., Pfeiffer, T., Fackler, O.T., Bosch, V., 2010a. Role of the C-terminal domain of the HIV-1 glycoprotein in cell-to-cell viral transmission between T lymphocytes. Retrovirology 7, 43.

Emerson, V., Holtkotte, D., Pfeiffer, T., Wang, I.H., Schnolzer, M., Kempf, T., Bosch, V., 2010b. Identification of the cellular prohibitin 1/prohibitin 2 heterodimer as an interaction partner of the C-terminal cytoplasmic domain of the HIV-1 glycoprotein. J. Virol. 84, 1355–1365.

Evans, D.T., Tillman, K.C., Desrosiers, R.C., 2002. Envelope glycoprotein cytoplasmic domains from diverse lentiviruses interact with the prenylated Rab acceptor J. Virol. 76, 327–337.

Freed, E.O., Martin, M.A., 1996. Domains of the human immunodeficiency virus type 1 matrix and gp41 cytoplasmic tail required for envelope incorporation into virions. J. Virol. 70, 341–351.

Green, N., Shinnick, T.M., Witte, O., Ponticelli, A., Sutcliffe, J.G., Lerner, R.A., 1981. Sequence-specific antibodies show that maturation of Moloney leukemia virus envelope polyprotein involves removal of a COOH-terminal peptide. Proc. Nat. Acad. Sci. U.S.A. 78, 6023–6027.

Haffar, O.K., Dowbenko, D.J., Berman, P.W., 1991. The cytoplasmic tail of HIV-1 gp160 contains regions that associate with cellular membranes. Virology 180, 439–441.

Henriksson, P., Bosch, V., 1998. Inhibition of cellular glycoprotein incorporation into human immunodeficiency virus-like particles by coexpression of additional cellular interaction partner. Virology 251, 16–21.

Kim, E.M., Lee, K.H., Kim, J.W., 1999. The cytoplasmic domain of HIV-1 gp41 interacts with the carboxyl-terminal region of alpha-catenin. Mol. Cells 9, 281–285.

Krausslich, H.G., Ochsenbauer, C., Traenckner, A.M., Mergener, K., Facke, M., Gelderblom, H.R., Bosch, V., 1993. Analysis of protein expression and virus-like particle formation in mammalian cell lines stably expressing HIV-1 gag and env gene products with or without active HIV proteinase. Virology 192, 605–617.

Li, M., Gao, F., Mascola, J.R., Stamatatos, L., Polonis, V.R., Koutsoukos, M., Voss, G., Goepfert, P., Gilbert, P., Greene, K.M., Bilska, M., Kothe, D.L., Salazar-Gonzalez, J. F., Wei, X., Decker, J.M., Hahn, B.H., Montefiori, D.C., 2005. Human immunodeficiency virus type 1 env clones from acute and early subtype B infections for standardized assessments of vaccine-elicited neutralizing antibodies. J. Virol. 79, 10108–10125.

Lodge, R., Gottlinger, H., Gabuzda, D., Cohen, E.A., Lemay, G., 1994. The intracytoplasmic domain of gp41 mediates polarized budding of human immunodeficiency virus type 1 in MDCK cells. J. Virol. 68, 4857–4861.

Lodge, R., Lalonde, J.P., Lemay, G., Cohen, E.A., 1997. The membrane-proximal intracytoplasmic tyrosine residue of HIV-1 envelope glycoprotein is critical for basolateral targeting of viral budding in MDCK cells. EMBO J. 16, 695–705.

Malim, M.H., Hauber, J., Fenrick, R., Cullen, B.R., 1988. Immunodeficiency virus rev trans-activator modulates the expression of the viral regulatory genes. Nature 335, 181–183.

Mermer, B., Felber, B.K., Campbell, M., Pavlakis, G.N., 1990. Identification of trans-dominant HIV-1 rev protein mutants by direct transfer of bacterially produced proteins into human cells. Nucleic Acids Res. 18, 2037–2044.

Miller, M.A., Mietzner, T.A., Cloyd, M.W., Robey, W.G., Montelaro, R.C., 1993. Identification of a calmodulin-binding and inhibitory peptide domain in the HIV-1 transmembrane glycoprotein. AIDS Res. Hum. Retroviruses 9, 1057–1066.

Murakami, T., Ablan, S., Freed, E.O., Tanaka, Y., 2004. Regulation of human immunodeficiency virus type 1 Env-mediated membrane fusion by viral protease activity. J. Virol. 78, 1026–1031.

Naldini, L., Blomer, U., Gallay, P., Ory, D., Mulligan, R., Gage, F.H., Verma, I.M., Trono, D., 1996. In vivo gene delivery and stable transduction of nondividing cells by a lentiviral vector. Science 272, 263–267.
Pohl, T., Zimmer, M., Mugele, K., Spiess, J., 1991. Primary structure and functional expression of a glutaminyl cyclase. Proc. Nat. Acad. Sci. U.S.A. 88, 10059–10063.
Postler, T.S., Desrosiers, R.C., 2012. The tale of the long tail: the cytoplasmic domain of HIV-1 gp41. J. Virol..
Ratner, L., Haseltine, W., Patarca, R., Livak, K.J., Starcich, B., Josephs, S.F., Doran, E.R., Rafalski, J.A., Whitehorn, E.A., Baumeister, K., et al., 1985. Complete nucleotide sequence of the AIDS virus, HTLV-III. Nature 313, 277–284.
Rein, A., Mirro, J., Haynes, J.G., Ernst, S.M., Nagashima, K., 1994. Function of the cytoplasmic domain of a retroviral transmembrane protein: p15E-p2E cleavage activates the membrane fusion capability of the murine leukemia virus Env protein. J. Virol. 68, 1773–1781.
Rice, N.R., Henderson, L.E., Sowder, R.C., Copeland, T.D., Oroszlan, S., Edwards, J.F., 1990. Synthesis and processing of the transmembrane envelope protein of equine infectious anemia virus. J. Virol. 64, 3770–3778.
Saido, T.C., Iwatsubo, T., Mann, D.M., Shimada, H., Ihara, Y., Kawashima, S., 1995. Dominant and differential deposition of distinct beta-amyloid peptide species, A beta N3(pE), in senile plaques. Neuron 14, 457–466.
Saido, T.C., Yamao-Harigaya, W., Iwatsubo, T., Kawashima, S., 1996. Amino- and carboxyl-terminal heterogeneity of beta-amyloid peptides deposited in human brain. Neurosci. Lett. 215, 173–176.
Schaal, H., Pfeiffer, P., Klein, M., Gehrmann, P., Scheid, A., 1993. Use of DNA end joining activity of a Xenopus laevis egg extract for construction of deletions and expression vectors for HIV-1 Tat and Rev proteins. Gene 124, 275–280.
Shacklett, B.L., Weber, C.J., Shaw, K.E., Keddie, E.M., Gardner, M.B., Sonigo, P., Luciw, P.A., 2000. The intracytoplasmic domain of the Env transmembrane protein is a locus for attenuation of simian immunodeficiency virus SIVmac in rhesus macaques. J. Virol. 74, 5836–5844.
Vale, W., Spiess, J., Rivier, C., Rivier, J., 1981. Characterization of a 41-residue ovine hypothalamic peptide that stimulates secretion of corticotropin and beta-endorphin. Science 213, 1394–1397.
Waheed, A.A., Ablan, S.D., Roser, J.D., Sowder, R.C., Schaffner, C.P., Chertova, E., Freed, E.O., 2007. HIV-1 escape from the entry-inhibiting effects of a cholesterol-binding compound via cleavage of gp41 by the viral protease. Proc. Nat. Acad. Sci. U.S.A. 104, 8467–8471.
Wilk, T., Pfeiffer, T., Bosch, V., 1992. Retained in vitro infectivity and cytopathogenicity of HIV-1 despite truncation of the C-terminal tail of the env gene product. Virology 189, 167–177.
Wilk, T., Pfeiffer, T., Bukovsky, A., Moldenhauer, G., Bosch, V., 1996. Glycoprotein incorporation and HIV-1 infectivity despite exchange of the gp160 membrane-spanning domain. Virology 218, 269–274.
Wyma, D.J., Jiang, J., Shi, J., Zhou, J., Lineberger, J.E., Miller, M.D., Aiken, C., 2004. Coupling of human immunodeficiency virus type 1 fusion to virion maturation: a novel role of the gp41 cytoplasmic tail. J. Virol. 78, 3429–3435.
Wyss, S., Berlioz-Torrent, C., Boge, M., Blot, G., Honing, S., Benarous, R., Thali, M., 2001. The highly conserved C-terminal dileucine motif in the cytosolic domain of the human immunodeficiency virus type 1 envelope glycoprotein is critical for its association with the AP-1 clathrin adaptor [correction of adapter]. J. Virol. 75, 2982–2992.
Zhang, H., Wang, L., Kao, S., Whitehead, I.P., Hart, M.J., Liu, B., Duus, K., Burridge, K., Der, C.J., Su, L., 1999. Functional interaction between the cytoplasmic leucine-zipper domain of HIV-1 gp41 and p115-RhoGEF. Curr. Biol. 9, 1271–1274.